



**THIS PAGE BLANK (USPTO)**

## 明 細 書

デジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体

### 技術分野

本発明はデジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体に関し、レートコンバータ又はPCM (Pulse Code Modulation) 復号装置等においてデジタル信号に対してデータの補間処理を行うデジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体に適用して好適なものである。

### 背景技術

従来、デジタルオーディオ信号をデジタル／アナログコンバータに入力する前に、サンプリング周波数を元の値の数倍に変換するオーバサンプリング処理を行っている。これにより、デジタル／アナログコンバータから出力されたデジタルオーディオ信号はアナログ・アンチ・エイリアス・フィルタの位相特性が可聴周波数高域で一定に保たれ、また、サンプリングに伴うデジタル系のイメージ雑音の影響が排除されるようになされている。

かかるオーバサンプリング処理では、通常、線形一次（直線）補間方式のデジタルフィルタが用いられている。このようなデジタルフィルタは、サンプリングレートが変わったりデータが欠落した場合等に、複数の既存データの平均値を求めて直線的な補間データを生成するものである。

ところが、オーバサンプリング処理後のデジタルオーディオ信号は、線形一次補間によって時間軸方向に対してデータ量が数倍に緻密になっているものの、オーバサンプリング処理後のデジタルオーディオ信号の周波数帯域は変換前とあまり変わらず、音質そのものは向上していない。さらに、補間されたデータは必ずしもA/D変換前のアナログオーディオ信号の波形に基づいて生成されたの

ではないため、波形再現性もほとんど向上していない。

また、サンプリング周波数の異なるデジタルオーディオ信号をダビングする場合において、サンプリング・レート・コンバータを用いて周波数を変換しているが、かかる場合でも線形一次デジタルフィルタによって直線的なデータの補間しか行うことができず、音質や波形再現性を向上することが困難であった。さらに、デジタルオーディオ信号のデータサンプルが欠落した場合において同様である。

#### 発明の開示

本発明は以上の点を考慮してなされたもので、デジタル信号の波形再現性を一段と向上し得るデジタル信号処理方法、学習方法及びそれらの装置並びにプログラム格納媒体を提案しようとするものである。

かかる課題を解決するため本発明においては、入力デジタル信号の包絡線に基づいて入力デジタル信号のクラスを分類し、当該分類されたクラスに対応した予測方式で入力デジタル信号を変換するようにしたことにより、一段と入力デジタル信号の特徴に適応した変換を行うことができる。

#### 図面の簡単な説明

図 1 は、本発明によるデジタル信号処理装置の第 1 の実施の形態を示すブロック図である。

図 2 は、包絡線を用いたクラス分類適応処理の説明に供する信号波形図である。

図 3 は、オーディオ信号処理装置の構成を示すブロック図である。

図 4 は、第 1 の実施の形態のオーディオ信号変換処理手順を示すフローチャートである。

図 5 は、包絡線の算出処理手順を示すフローチャートである。

図 6 は、包絡線の算出方法の説明に供する信号波形図である。

図 7 は、包絡線の算出方法の説明に供する信号波形図である。

図 8 は、包絡線の算出方法の説明に供する信号波形図である。

図 9 は、包絡線の算出方法の説明に供する信号波形図である。

図 10 は、包絡線の算出方法の説明に供する信号波形図である。

図 11 は、本発明による学習装置の第 1 の実施の形態を示すブロック図である。

図 12 は、デジタル信号処理装置の他の実施の形態を示すブロック図である。

図 13 は、学習装置の他の実施の形態を示すブロック図である。

図 14 は、本発明によるデジタル信号処理装置の第 2 の実施の形態を示すブロック図である。

図 15 は、第 2 の実施の形態によるクラス分類適応処理の説明に供する信号波形図である。

図 16 は、第 2 の実施の形態のオーディオ信号変換処理手順を示すフローチャートである。

図 17 は、本発明による学習装置の第 2 の実施の形態を示すブロック図である。

## 発明を実施するための最良の形態

以下図面について、本発明の一実施の形態を詳述する。

### (1) 第 1 の実施の形態

図 1 においてオーディオ信号処理装置 10 は、デジタルオーディオ信号（以下これをオーディオデータと呼ぶ）のサンプリングレートを上げたり、オーディオデータを補間する際に、真値に近いオーディオデータをクラス分類適用処理によって生成するようになされている。因みに、デジタルオーディオ信号とは、人や動物が発する声を表す音声信号、楽器が発する楽音を表す楽音信号、及びその他の音を表す信号を意味するものである。

すなわち、オーディオ信号処理装置 10 において、包絡線算出部 11 は入力端子  $T_{IN}$  から供給された図 2 (A) に示す入力オーディオデータ D10 を所定時間毎の領域（この実施の形態の場合、例えば 6 サンプル毎とする）に分割した後、当該分割された各時間領域の波形について、後述する包絡線算出方法によりその包絡線を算出する。

包絡線算出部 11 は入力オーディオデータ D10 のこのとき分割された時間領域の包絡線算出結果を入力オーディオデータ D10 の包絡線波形データ D11（図 2 (B)）としてクラス分類部 14 に供給する。

また、クラス分類部抽出部 12 は入力端子  $T_{IN}$  から供給された図 2 (A) に示す入力オーディオデータ D10 を、包絡線算出部 11 の場合と同様の時間領域（この実施の形態の場合例えば 6 サンプル）に分割することによりクラス分類しようとするオーディオ波形データ D12 を抽出し、これをクラス分類部 14 に供給する。

クラス分類部 14 は、クラス分類抽出部 12 において切り出されたオーディオ波形データ D12 に対応する包絡線波形データ D11 について、当該包絡線波形データ D11 を圧縮して圧縮データパターンを生成する ADRC (Adaptive Dynamic Range Coding) 回路部と、包絡線波形データ D11 の属するクラスコードを発生するクラスコード発生回路部とを有する。

ADRC 回路部は包絡線波形データ D11 に対して、例えば 8 ビットから 2 ビットに圧縮するような演算を行うことによりパターン圧縮データを形成する。この ADRC 回路部は、適応的量子化を行うものであり、ここでは、信号レベルの局所的なパターンを短い語長で効率的に表現することができるので、信号パターンのクラス分類のコード発生用に用いられる。

具体的には、包絡線波形上の 6 つの 8 ビットのデータ（包絡線波形データ）をクラス分類しようとする場合、 $2^48$  という膨大な数のクラスに分類しなければならず、回路上の負担が多くなる。そこで、この実施の形態のクラス分類部 14

ではその内部に設けられたA D R C回路部で生成されるパターン圧縮データに基づいてクラス分類を行う。例えば6つの包絡線波形データに対して1ビットの量子化を実行すると、6つの包絡線波形データを6ビットで表すことができ、 $2^6 = 64$ クラスに分類することができる。

ここで、A D R C回路部は、切り出された領域内の包絡線のダイナミックレンジをD R、ビット割り当をm、各包絡線波形データのデータレベルをL、量子化コードをQとすると、次式、

$$D R = M A X - M I N + 1$$

$$Q = \{(L - M I N + 0.5) \times 2^m / D R\} \dots\dots (1)$$

に従って、領域内の最大値M A Xと最小値M I Nとの間を指定されたビット長で均等に分割して量子化を行う。なお、(1)式において{ }は小数点以下の切り捨て処理を意味する。かくして包絡線算出部11において算出された包絡線上の6つの波形データが、それぞれ例えば8ビット( $m=8$ )で構成されているとすると、これらはA D R C回路部においてそれぞれが2ビットに圧縮される。

このようにして圧縮された包絡線波形データをそれぞれ $q_n$  ( $n=1 \sim 6$ )とすると、クラス分類部14に設けられたクラスコード発生回路部は、圧縮された包絡線波形データ $q_n$ に基づいて、次式、

$$c l a s s = \sum_{i=1}^n q_i (2^P)^i \dots\dots (2)$$

に示す演算を実行することにより、そのブロック( $q_1 \sim q_6$ )が属するクラスを示すクラスコードc l a s sを算出し、当該算出されたクラスコードc l a s

sを表すクラスコードデータD14を予測係数メモリ15に供給する。このクラスコードclassは、予測係数メモリ15から予測係数を読み出す際の読み出しアドレスを示す。因みに(2)式において、nは圧縮された包絡線波形データ $q_n$ の数を表し、この実施の形態の場合 $n=6$ であり、またPはビット割り当てを表し、この実施の形態の場合 $P=2$ である。

このようにして、クラス分類部14はクラス分類部抽出部12において入力オーディオデータD10から切り出されたオーディオ波形データD12に対応する包絡線波形データD11のクラスコードデータD14を生成し、これを予測係数メモリ15に供給する。

予測係数メモリ15には、各クラスコードに対応する予測係数のセットがクラスコードに対応するアドレスにそれぞれ記憶されており、クラス分類部14から供給されるクラスコードデータD14に基づいて、当該クラスコードに対応するアドレスに記憶されている予測係数のセット $w_1 \sim w_n$ が読み出され、予測演算部16に供給される。

予測演算部16は、予測演算部抽出部13において入力オーディオデータD10から時間軸領域で切り出された予測演算しようとするオーディオ波形データ(予測タップ)D13( $x_1 \sim x_n$ )と、予測係数 $w_1 \sim w_n$ に対して、次式

$$y' = w_1 x_1 + w_2 x_2 + \dots + w_n x_n \quad \dots\dots (3)$$

に示す積和演算を行うことにより、予測結果 $y'$ を得る。この予測値 $y'$ が、音質が改善されたオーディオデータD16(図2(C))として予測演算部16から出力される。

なお、オーディオ信号処理装置10の構成として図1について上述した機能ブロックを示したが、この機能ブロックを構成する具体的構成として、この実施の形態においては図3に示すコンピュータ構成の装置を用いる。すなわち、図3において、オーディオ信号処理装置10は、バスBUSを介してCPU21、RO



M (Read Only Memory) 22、予測係数メモリ 15 を構成する RAM (Random Access Memory) 15、及び各回路部がそれぞれ接続された構成を有し、CPU 11 は ROM 22 に格納されている種々のプログラムを実行することにより、図 1 について上述した各機能ブロック（包絡線算出部 11、クラス分類部抽出部 12、予測演算部抽出部 13、クラス分類部 14 及び予測演算部 16）として動作するようになされている。

また、オーディオ信号処理装置 10 にはネットワークとの間で通信を行う通信インターフェース 24、フロッピディスクや光磁気ディスク等の外部記憶媒体から情報を読み出すリムーバブルドライブ 28 を有し、ネットワーク経由又は外部記憶媒体から図 1 について上述したクラス分類適用処理を行うための各プログラムをハードディスク装置 25 のハードディスクに読み込み、当該読み込まれたプログラムに従ってクラス分類適用処理を行うこともできる。

ユーザは、キーボードやマウス等の入力手段 26 を介して種々のコマンドを入力することにより、CPU 21 に対して図 1 について上述したクラス分類処理を実行させる。この場合、オーディオ信号処理装置 10 はデータ入出力部 27 を介して音質を向上させようとするオーディオデータ（入力オーディオデータ）D10 を入力し、当該入力オーディオデータ D10 に対してクラス分類適用処理を施した後、音質が向上したオーディオデータ D16 をデータ入出力部 27 を介して外部に出力し得るようになされている。

因みに、図 4 はオーディオ信号処理装置 10 におけるクラス分類適用処理の処理手順を示し、オーディオ信号処理装置 10 はステップ SP101 から当該処理手順に入ると、続くステップ SP102 において入力オーディオデータ D10 の包絡線を包絡線算出部 11 において算出する。

この算出された包絡線は入力オーディオデータ D10 の特徴を表すものであり、オーディオ信号処理装置 10 は、ステップ SP103 に移ってクラス分類部 14 により包絡線に基づいてクラスを分類する。そしてオーディオ信号処理装置 10 はクラス分類の結果得られたクラスコードを用いて予測係数メモリ 15 から予

測係数を読み出す。この予測係数は予め学習によりクラス毎に対応して格納されており、オーディオ信号処理装置10はクラスコードに対応した予測係数を読み出すことにより、このときの包絡線の特徴に合致した予測係数を用いることができる。

予測係数メモリ15から読み出された予測係数は、ステップSP104において予測演算部16の予測演算に用いられる。これにより、入力オーディオデータD10はその包絡線の特徴に適応した予測演算により、所望とするオーディオデータD16に変換される。かくして入力オーディオデータD10はその音質が改善されたオーディオデータD16に変換され、オーディオ信号処理装置10はステップSP105に移って当該処理手順を終了する。

次に、オーディオ信号処理装置10の包絡線算出部11における入力オーディオデータD10の包絡線の算出方法について説明する。

すなわち、図5に示すように、包絡線算出部11（図1）は包絡線算出処理手順RT1に入ると、ステップSP1において外部から入力されてくる正負の極性がある入力オーディオデータD10をデータ入出力部27を介して入力し、続くステップSP2及びSP10に移る。

ステップSP2において包絡線算出部11は、図6に示すように、外部から入力されてくる正負の極性がある入力オーディオデータD10のうち、正領域AR1の信号成分のみを検出及び保持し、負領域AR2の信号線分をゼロレベルとし、ステップSP3に移る。

ステップSP3において包絡線算出部11は、図7に示すように、正領域AR1の入力オーディオデータD10の振幅がゼロレベルと重なるサンプリング時間位置DO1からその次に振幅がゼロレベルと重なるサンプリング時間位置DO2までの間（以下、これをゼロクロス間と呼ぶ）CR1での振幅の最大値 $x_1$ を検出し、当該最大値 $x_1$ が予め包絡線検出プログラムにて設定されている閾値よりも高い値であるか否かを判断する。

因みに、包絡線検出プログラムで予め設定される閾値は、ゼロクロス間の振幅

の最大値  $x_1$  を包絡線の候補値（サンプリング点）とするか否かを決定づける所定の値となっており、結果としてなめらかな包絡線を検出し得るような値に設定しておき、このときの判断対象であるゼロクロス間  $CR_1$  の振幅の最大値  $x_1$  が当該閾値よりも高い値となる場合にはステップ  $SP_4$  に移る。また、このときの判断対象であるゼロクロス間の振幅の最大値が閾値よりも低い値となる場合、包絡線算出部 11 は、当該閾値よりも高い値となる最大値  $x_1$ （候補値（サンプリング点））が存在するゼロクロス間  $CR_1$  を検出するまで続ける。

ステップ  $SP_4$  において包絡線算出部 11 は、候補値（サンプリング点）とされた最大値  $x_1$  が存在するゼロクロス間  $CR_1$  の次のゼロクロス間  $CR_2$  の最大値  $x_2$  を検出（図 7）し、ステップ  $SP_5$  に移る。

ステップ  $SP_5$  において包絡線算出部 11 は、ステップ  $SP_3$  及び  $SP_4$  において得られた各最大値  $x_1$  及び  $x_2$  に対して  $f(t) = p(t_2 - t_1)$  で表される関数によって算出された値に最大値  $x_1$  を乗じた値が当該最大値  $x_2$  よりも高い値であるか否かを判断する。

因みに、関数  $f(t)$  において、「 $t_2$ 」及び「 $t_1$ 」は、各最大値  $x_1$  及び  $x_2$  が検出されたサンプリング時間位置を表しており、例えばこのとき入力される信号（入力オーディオデータ  $D10$ ）がサンプリング周波数  $8\text{ kHz}$ 、量子化  $16\text{ bit}$  と想定した場合、ゼロクロス間のサンプル数は  $5 \sim 20$  サンプルとなる場合が多いため、「 $t_2$ 」及び「 $t_1$ 」においても  $5 \sim 20$  サンプルとなる。また、「 $p$ 」は任意に設定し得るパラメータであり、例えばこのとき入力される信号（入力オーディオデータ  $D10$ ）がサンプリング周波数  $8\text{ kHz}$ 、量子化  $16\text{ bit}$  と想定した場合、 $p = -90$  などとされる。

さらに関数  $f(t) = p(t_2 - t_1)$  で表される値に最大値  $x_1$  を乗じた値は、最大値  $x_1$  及び  $x_2$  間の傾きを表すようになされており、当該関数  $f(t) = p(t_2 - t_1)$  で表される値に最大値  $x_1$  を乗じた値よりも最大値  $x_2$  の値のほうが大きい場合には、最大値  $x_1$  と最大値  $x_2$  との振幅差が少ないことにより、結果としてなめらかな包絡線を検出し得る。従って、このときの判断対象で

ある最大値  $x_2$  が当該関数によって表される値に最大値  $x_1$  を乗じた値よりも高い値となる場合にはステップ S P 5 において肯定結果を得、続くステップ S P 6 に移る。

これに対して最大値  $x_2$  が当該関数によって表される値に最大値  $x_1$  を乗じた値よりも低い値となる場合には、ステップ S P 4 において、関数によって表される値に最大値  $x_1$  を乗じた値よりも高い値である最大値  $x_2$  (図 7) が検出されるまでゼロクロス間 ( $CR_3 \cdots CR_n$ ) の振幅の最大値  $x_2$  (図 7) を検出するようになされており、このとき再度検出して得られた最大値  $x_2$  と、ステップ S P 3 において得られた最大値  $x_1$  とに対して、 $f(t) = p(t_2 - t_1)$  で表される関数によって算出された値に最大値  $x_1$  を乗じた値が再度検出して得られた最大値  $x_2$  よりも高い値であると判断されるまで最大値  $x_2$  の検出を繰り返す。

ステップ S P 6 において包絡線算出部 11 は、包絡線の候補値 (サンプリング点) とされた最大値  $x_1$  及び最大値  $x_2$  間のデータに対して線形一次補間方法を用いて補間処理を施し、続くステップ S P 7 及び S P 8 に移る。

ステップ S P 7 において包絡線算出部 11 は、補間処理を施した最大値  $x_1$  及び最大値  $x_2$  間のデータ及び候補値 (サンプリング点) を包絡線データ D 11 (図 1) として、クラス分類部 14 (図 1) に出力する。

また、ステップ S P 8 において包絡線算出部 11 は、外部から入力されてくる入力オーディオデータ D 10 が全て入力されたか否かを判断する。ここで否定結果が得られると、このことは入力オーディオデータ D 10 が続いて入力されていることを表しており、このとき包絡線算出部 11 は、ステップ S P 3 に戻って、入力オーディオデータ D 10 の正領域 A R 1 からゼロクロス間 C R 1 の振幅の最大値  $x_1$  を再び検出する。

これに対して、ステップ S P 8 において肯定結果が得られると、このことは入力オーディオデータ D 10 を全て入力し終わったことを表しており、このとき包絡線算出部 11 は、ステップ S P 20 に移って包絡線算出処理手順 R T 1 を終了

する。

一方、ステップSP10において包絡線算出部11は、外部から入力される正負の極性がある入力オーディオデータD10のうち、負領域AR2（図6）の信号成分のみを検出及び保持し、正領域AR1（図6）の信号成分をゼロレベルとし、ステップSP11に移る。

ステップSP11において包絡線算出部11は、図8に示すように、負領域AR2のゼロクロス間CR11の振幅の最大値 $x_{11}$ を検出し、ステップSP3と同様に最大値 $x_{11}$ が予め包絡線検出プログラムにて設定されている閾値よりも負方向に高い値であるか否かを判断する。ここで肯定結果が得られた（すなわち、閾値よりも負方向に高い値である）場合には、ステップSP12に移り、否定結果が得られた（すなわち、閾値よりも負方向に低い値である）場合には、閾値よりも負方向に高い値となる最大値 $y_{11}$ が検出されるまでステップSP11の検出処理を続ける。

ステップSP12において包絡線算出部11は、候補値（サンプリング点）とされた最大値 $x_{11}$ が含まれるゼロクロス間CR'1の次のゼロクロス間CR'2の振幅の最大値 $x_{12}$ を検出（図8）し、ステップSP13に移る。

ステップSP13において包絡線算出部11は、ステップSP5と同様にステップSP11及びSP12において得られた各最大値 $x_{11}$ 及び $x_{12}$ に対して $f(t) = p(t_{12} - t_{11})$ で表される関数によって算出された値に最大値 $x_{11}$ を乗じた値が最大値 $x_{12}$ よりも負の方向に高い値であるか否かを判断する。因みに、「 $p$ 」は任意に設定し得るパラメータであり、例えばこのとき入力される入力オーディオデータD10がサンプリング周波数8kHz、量子化16bitと想定した場合、 $p = 90$ などとされる。

包絡線算出部11は、ステップSP13において、肯定結果が得られた（すなわち、 $f(t) = p(t_{12} - t_{11})$ で表される関数によって算出された値に最大値 $x_{11}$ を乗じた値が最大値 $x_{12}$ よりも負の方向に高い値である）場合には、ステップSP14に移り、否定結果が得られた（すなわち、 $f(t) = p(t$

$t_{12} - t_{11}$ ) で表される関数によって算出された値に最大値  $x_{11}$  を乗じた値が最大値  $x_{12}$  よりも負の方向に低い値である) 場合には、ステップ SP 1 2 において、 $f(t) = p(t_{12} - t_{11})$  で表される関数によって算出された値に最大値  $x_{11}$  を乗じた値よりも負の方向に高い値である最大値  $x_{12}$  (図 8) が検出されるまでゼロクロス間 ( $CR'_3 \cdots CR'_n$ ) の振幅の最大値  $x_{12}$  (図 8) を検出する。

ステップ SP 1 4 において包絡線算出部 1 1 は、包絡線の候補値 (サンプリング点) とされた最大値  $x_{11}$  及び最大値  $x_{12}$  間のデータに対して線形一次補間方法を用いて補間処理を施し、続くステップ SP 7 及び SP 1 5 に移る。

ステップ SP 7 において包絡線算出部 1 1 は、補間処理を施した最大値  $x_{11}$  及び最大値  $x_{12}$  間のデータ及び候補値 (サンプリング点) を包絡線データ D 1 1 (図 1) として、クラス分類部 1 4 (図 1) に出力する。

また、ステップ SP 1 5 において包絡線算出部 1 1 は、外部から入力されてくる入力オーディオデータ D 1 0 が全て入力されたか否かを判断する。ここで否定結果が得られると、このことは入力オーディオデータ D 1 0 が続いて入力されていることを表しており、このとき包絡線算出部 1 1 は、ステップ SP 1 1 に戻って、入力オーディオデータ D 1 0 の負領域 AR 2 からゼロクロス間の振幅の最大値  $x_{11}$  を再び検出する。

これに対して、ステップ SP 1 5 において肯定結果が得られると、このことは入力オーディオデータ D 1 0 を全て入力し終わったことを表しており、このとき包絡線算出部 1 1 は、ステップ SP 2 0 に移って包絡線算出処理手順 RT 1 を終了する。

このように、包絡線算出部 1 1 は簡単な包絡線算出アルゴリズムによって、結果として正領域 AR 1 では図 9 に示すような、なめらかな包絡線 ENV 5、及び負領域 AR 2 では図 10 に示すような、なめらかな包絡線 ENV 6 を生成させ得る包絡線データ (候補値 (サンプリング点) と、補間処理を行った候補値間のデータ) をリアルタイムに算出することができる。

次に、図 1 について上述した予測係数メモリ 15 に記憶するクラス毎の予測係数のセットを予め学習によって得るための学習回路について説明する。

図 11 において、学習回路 30 は、高音質の教師オーディオデータ D30 を生徒信号生成フィルタ 37 に受ける。生徒信号生成フィルタ 37 は、間引き率設定信号 D39 により設定された間引き率で教師オーディオデータ D30 を所定時間ごとに所定サンプル間引くようになされている。

この場合、生徒信号生成フィルタ 37 における間引き率によって、生成される予測係数が異なり、これに応じて上述のオーディオ信号処理装置 10 で再現されるオーディオデータも異なる。例えば、上述のオーディオ信号処理装置 10 においてサンプリング周波数を高くすることでオーディオデータの音質を向上しようとする場合、生徒信号生成フィルタ 37 ではサンプリング周波数を減らす間引き処理を行う。また、これに対して上述のオーディオ信号処理装置 10 において入力オーディオデータ D10 の欠落したデータサンプルを補うことで音質の向上を図る場合には、これに応じて、生徒信号生成フィルタ 37 ではデータサンプルを欠落させる間引き処理を行うようになされている。

かくして、生徒信号生成フィルタ 37 は教師オーディオデータ 30 から所定の間引き処理により生徒オーディオデータ D37 を生成し、これを包絡線算出部 31、クラス分類部抽出部 32 及び予測演算部抽出部 33 にそれぞれ供給する。

包絡線算出部 31 は生徒信号生成フィルタ 37 から供給された生徒オーディオデータ D37 を所定時間毎の領域（この実施の形態の場合、例えば 6 サンプル毎とする）に分割した後、当該分割された各時間領域の波形について、図 5 について上述した包絡線算出方法によりその包絡線を算出する。

包絡線算出部 31 は生徒オーディオデータ D37 のこのとき分割された時間領域の包絡線算出結果を生徒オーディオデータ D37 の包絡線波形データ D31 としてクラス分類部 34 に供給する。

また、クラス分類部抽出部 32 は生徒信号生成フィルタ 37 から供給された生徒オーディオデータ D37 を、包絡線算出部 31 の場合と同様の時間領域（この

実施の形態の場合例えば6サンプル)に分割することによりクラス分類しようとするオーディオ波形データD32を抽出し、これをクラス分類部34に供給する。

クラス分類部34は、クラス分類抽出部32において切り出されたオーディオ波形データD32に対応する包絡線波形データD31について、当該包絡線波形データD31を圧縮して圧縮データパターンを生成するADRC (Adaptive Dynamic Range Coding) 回路部と、包絡線波形データD31の属するクラスコードを発生するクラスコード発生回路部とを有する。

ADRC回路部は包絡線波形データD31に対して、例えば8ビットから2ビットに圧縮するような演算を行うことによりパターン圧縮データを形成する。このADRC回路部は、適応的量子化を行うものであり、ここでは、信号レベルの局所的なパターンを短い語長で効率的に表現することができるので、信号パターンのクラス分類のコード発生用に用いられる。

具体的には、包絡線波形上の6つの8ビットのデータ(包絡線波形データ)をクラス分類しようとする場合、 $2^48$ という膨大な数のクラスに分類しなければならず、回路上の負担が多くなる。そこで、この実施の形態のクラス分類部14ではその内部に設けられたADRC回路部で生成されるパターン圧縮データに基づいてクラス分類を行う。例えば6つの包絡線波形データに対して1ビットの量子化を実行すると、6つの包絡線波形データを6ビットで表すことができ、 $2^6 = 64$ クラスに分類することができる。

ここで、ADRC回路部は、切り出された領域内の包絡線のダイナミックレンジをDR、ビット割り当をm、各包絡線波形データのデータレベルをL、量子化コードをQとして、上述の(1)式と同様の演算により、領域内の最大値MAXと最小値MINとの間を指定されたビット長で均等に分割して量子化を行う。かくして包絡線算出部1において算出された包絡線上の6つの波形データが、それぞれ例えば8ビット( $m=8$ )で構成されているとすると、これらはADRC回



路部においてそれぞれが2ビットに圧縮される。

このようにして圧縮された包絡線波形データをそれぞれ $q_n$  ( $n=1\sim6$ )とすると、クラス分類部34に設けられたクラスコード発生回路部は、圧縮された包絡線波形データ $q_n$ に基づいて、上述の(2)式と同様の演算を実行することにより、そのブロック( $q_1\sim q_6$ )が属するクラスを示すクラスコード $class$ を算出し、当該算出されたクラスコード $class$ を表すクラスコードデータD34を予測係数算出部36に供給する。因みに(2)式において、 $n$ は圧縮された包絡線波形データ $q_n$ の数を表し、この実施の形態の場合 $n=6$ であり、また $P$ はビット割り当てを表し、この実施の形態の場合 $P=2$ である。

このようにして、クラス分類部34はクラス分類部抽出部32において切り出されたオーディオ波形データD32に対応する包絡線波形データD31のクラスコードデータD34を生成し、これを予測係数算出部36に供給する。また、予測係数算出部36には、クラスコードデータD34に対応した時間軸領域のオーディオ波形データD33( $x_1, x_2, \dots, x_n$ )が予測演算部抽出部33において切り出されて供給される。

予測係数算出部36は、クラス分類部34から供給されたクラスコード $class$ と、各クラスコード $class$ 毎に切り出されたオーディオ波形データD33と、入力端 $T_{IN}$ から供給された高音質の教師オーディオデータD30とを用いて、正規方程式を立てる。

すなわち、生徒オーディオデータD37の $n$ サンプルのレベルをそれぞれ $x_1, x_2, \dots, x_n$ として、それぞれに $p$ ビットのADRCを行った結果の量子化データを $q_1, \dots, q_n$ とする。このとき、この領域のクラスコード $class$ を上述の(2)式のように定義する。そして、上述のように生徒オーディオデータD37のレベルをそれぞれ、 $x_1, x_2, \dots, x_n$ とし、高音質の教師オーディオデータD30のレベルを $y$ としたとき、クラスコード毎に、予測係数 $w_1, w_2, \dots, w_n$ による $n$ タップの線形推定式を設定する。これを次式、

$$y = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n \quad \cdots \cdots (4)$$

とする。学習前は、 $w_n$ が未定係数である。

学習回路30では、クラスコード毎に、複数のオーディオデータに対して学習を行う。データサンプル数がMの場合、上述の(4)式に従って、次式、

$$y_k = w_1 x_{k1} + w_2 x_{k2} + \cdots + w_n x_{kn} \quad \cdots \cdots (5)$$

が設定される。但し  $k = 1, 2, \cdots, M$  である。

$M > n$  の場合、予測係数  $w_1, \cdots, w_n$  は一意的に決まらないので、誤差ベクトル  $e$  の要素を次式、

$$e_k = y_k - \{w_1 x_{k1} + w_2 x_{k2} + \cdots + w_n x_{kn}\} \quad \cdots \cdots (6)$$

によって定義し (但し、 $k = 1, 2, \cdots, M$ )、次式、

$$e^2 = \sum_{k=0}^M e_k^2 \quad \cdots \cdots (7)$$

を最小にする予測係数を求める。いわゆる、最小自乗法による解法である。

ここで、(7)式による  $w_n$  の偏微分係数を求める。この場合、次式、

$$\begin{aligned} \frac{\partial e^2}{\partial w_i} &= \sum_{k=0}^M 2 \left[ \frac{\partial e_k}{\partial w_i} \right] e_k = \sum_{k=0}^M 2 X_{ki} \cdot e_k \\ &= \sum_{k=0}^M 2 X_{ki} \cdot e_k \quad (i = 1, 2, \cdots, n) \end{aligned} \quad \cdots \cdots (8)$$

を「0」にするように、各 $w_n$  ( $n=1\sim6$ )を求めれば良い。

そして、次式、

$$X_{ij} = \sum_{p=0}^M X_{pi} \cdot X_{pj} \quad \dots\dots (9)$$

$$Y_i = \sum_{k=0}^M X_{ki} \cdot y_k \quad \dots\dots (10)$$

のように、 $X_{ij}$ 、 $Y_i$  を定義すると、(8)式は行列を用いて次式、

$$\begin{bmatrix} X_{11} & X_{12} & \dots\dots\dots X_{1n} \\ X_{21} & X_{22} & \dots\dots\dots X_{2n} \\ \vdots & \vdots & \\ \vdots & \vdots & \\ X_{m1} & X_{m2} & \dots\dots\dots X_{mn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ \vdots \\ Y_n \end{bmatrix} \quad \dots\dots (11)$$

として表される。

この方程式は、一般に正規方程式と呼ばれている。なお、ここでは $n=6$ である。

全ての学習用データ（教師オーディオデータD30、クラスコードclass、オーディオ波形データD33）の入力が完了した後、予測係数算出部36は各クラスコードclassに上述の(11)式に示した正規方程式を立てて、この正規方程式を掃き出し法等の一般的な行列解法を用いて、各 $W_n$ について解き、

各クラスコード毎に、予測係数を算出する。予測係数算出部 36 は、算出された各予測係数 (D 36) を予測係数メモリ 15 に書き込む。

このような学習を行った結果、予測係数メモリ 15 には、量子化データ  $q_1$ 、……、 $q_6$  で規定されるパターン毎に、高音質のオーディオデータ  $y$  を推定するための予測係数が、各クラスコード毎に格納される。この予測係数メモリ 15 は、図 1 について上述したオーディオ信号処理装置 10 において用いられる。かかる処理により、線形推定式に従って通常のオーディオデータから高音質のオーディオデータを作成するための予測係数の学習が終了する。

このように、学習回路 30 は、オーディオ信号処理装置 10 において補間処理を行う程度を考慮して、生徒信号生成フィルタ 37 で高音質の教師オーディオデータの間引き処理を行うことにより、オーディオ信号処理装置 10 における補間処理のための予測係数を生成することができる。

以上の構成において、オーディオ信号処理装置 10 は、包絡線算出部 11 において入力オーディオデータ D 10 の時間波形領域での包絡線を算出する。この包絡線は入力オーディオデータ D 10 の音質ごとに変わるもので、オーディオ信号処理装置 10 は入力オーディオデータ D 10 の包絡線に基づいてそのクラスを特定する。

オーディオ信号処理装置 10 は、予め学習時に例えば歪みのない高音質のオーディオデータ (教師オーディオデータ) を得るための予測係数をクラス毎に求めておき、包絡線に基づいてクラス分類された入力オーディオデータ D 10 をそのクラスに応じた予測係数により予測演算する。これにより、入力オーディオデータ D 10 はその音質に応じた予測係数を用いて予測演算されるので、実用上十分な程度に音質が向上する。

また、クラス毎の予測係数を生成する学習時において、位相の異なる多数の教師オーディオデータについてそれぞれに対応した予測係数を求めておくことにより、オーディオ信号処理装置 10 における入力オーディオデータ D 10 のクラス分類適応処理時に位相変動が生じて、位相変動に対応した処理を行うことがで

きる。

以上の構成によれば、入力オーディオデータD10の時間波形領域における包絡線に基づいて入力オーディオデータD10をクラス分類し、当該クラス分類された結果に基づく予測係数を用いて入力オーディオデータD10を予測演算するようにしたことにより、入力オーディオデータD10を一段と高音質のオーディオデータD16に変換することができる。

なお上述の実施の形態においては、オーディオ信号処理装置10及び学習装置30において、クラス分類部抽出部12、32及び予測演算部抽出部13、33により入力オーディオデータD10、D37を常に一定の範囲毎に切り出す場合について述べたが、本発明はこれに限らず、例えば図1及び図11との対応部分に同一符号を付して示す図12及び図13に示すように、包絡線算出部11、31において算出された包絡線の特徴に基づいて抽出制御信号CONT11、CONT31を可変クラス分類部抽出部12'、可変予測演算部抽出部13'及び可変クラス分類部抽出部32'、可変予測演算部抽出部33'に供給することにより入力オーディオデータD10、D37の切り出し範囲を制御するようにしても良い。

また上述の実施の形態においては、包絡線データD11に基づいてクラス分類する場合について述べたが、本発明はこれに限らず、クラス分類部抽出部12において入力オーディオデータD10の波形からクラス分類を行うと共に、包絡線算出部11において包絡線のクラスを算出し、これら2つのクラス情報をクラス分類部14において統合することにより、入力オーディオデータD10の波形とその包絡線との両方に基づくクラス分類を行うようにしても良い。

## (2) 第2の実施の形態

図1との対応部分に同一符号を付して示す図14において包絡線算出部11は入力端子T<sub>IN</sub>から供給された図15(A)に示す入力オーディオデータD10を所定時間毎の領域（この実施の形態の場合、例えば6サンプル毎とする）に分割した後、当該分割された各時間領域の波形について、図5について上述した包

絡線算出方法によりその包絡線を算出する。

包絡線算出部11は入力オーディオデータD10のこのとき分割された時間領域の包絡線算出結果を入力オーディオデータD10の包絡線波形データD11（図15（C））としてクラス分類部14、包絡線残差算出部111及び包絡線予測演算部116に供給する。

包絡線残差算出部111は入力オーディオデータD10と、包絡線算出部11から供給された包絡線データD11との残差を求め、これを正規化部112において正規化することにより、入力オーディオデータD10の搬送波D112（図15（B））を抽出し、これを変調部117に供給する。

クラス分類部14は、包絡線波形データD11について、当該包絡線波形データD11を圧縮して圧縮データパターンを生成するADRC（Adaptive Dynamic Range Coding）回路部と、包絡線波形データD11の属するクラスコードを発生するクラスコード発生回路部とを有する。

ADRC回路部は包絡線波形データD11に対して、例えば8ビットから2ビットに圧縮するような演算を行うことによりパターン圧縮データを形成する。このADRC回路部は、適応的量子化を行うものであり、ここでは、信号レベルの局所的なパターンを短い語長で効率的に表現することができるので、信号パターンのクラス分類のコード発生用に用いられる。

具体的には、包絡線波形上の6つの8ビットのデータ（包絡線波形データ）をクラス分類しようとする場合、 $2^48$ という膨大な数のクラスに分類しなければならず、回路上の負担が多くなる。そこで、この実施の形態のクラス分類部14ではその内部に設けられたADRC回路部で生成されるパターン圧縮データに基づいてクラス分類を行う。例えば6つの包絡線波形データに対して1ビットの量子化を実行すると、6つの包絡線波形データを6ビットで表すことができ、 $2^6 = 64$ クラスに分類することができる。

ここで、ADRC回路部は、切り出された領域内の包絡線のダイナミックレンジをDR、ビット割り当をm、各包絡線波形データのデータレベルをL、量子化

コードを $Q$ とすると、上述の(1)式に従って、領域内の最大値 $MAX$ と最小値 $MIN$ との間を指定されたビット長で均等に分割して量子化を行う。なお、(1)式において $\{ \}$ は小数点以下の切り捨て処理を意味する。かくして包絡線算出部1において算出された包絡線上の6つの波形データが、それぞれ例えば8ビット( $m=8$ )で構成されているとすると、これらは $ADRC$ 回路部においてそれぞれが2ビットに圧縮される。

このようにして圧縮された包絡線波形データをそれぞれ $q_n$  ( $n=1\sim 6$ )とすると、クラス分類部14に設けられたクラスコード発生回路部は、圧縮された包絡線波形データ $q_n$ に基づいて、上述の(2)式に示す演算を実行することにより、そのブロック( $q_1\sim q_6$ )が属するクラスを示すクラスコード $class$ を算出し、当該算出されたクラスコード $class$ を表すクラスコードデータ $D14$ を予測係数メモリ15に供給する。このクラスコード $class$ は、予測係数メモリ15から予測係数を読み出す際の読み出しアドレスを示す。

このようにして、クラス分類部14は包絡線波形データ $D11$ のクラスコードデータ $D14$ を生成し、これを予測係数メモリ15に供給する。

予測係数メモリ15には、各クラスコードに対応する予測係数のセットがクラスコードに対応するアドレスにそれぞれ記憶されており、クラス分類部14から供給されるクラスコードデータ $D14$ に基づいて、当該クラスコードに対応するアドレスに記憶されている予測係数のセット $W_1\sim W_n$ が読み出され、包絡線予測演算部116に供給される。

包絡線予測演算部116は、包絡線算出部11において算出された包絡線波形データ $D11$  ( $X_1\sim X_n$ )と、予測係数 $W_1\sim W_n$ に対して、上述の(3)式に示す積和演算を行うことにより、予測結果 $y'$ を得る。この予測値 $y'$ が、音質が改善されたオーディオデータの包絡線データ $D116$  (図14 (C))として変調部117に供給される。

かくして、変調部117は、包絡線残差算出部111から供給された搬送波 $D112$ を包絡線データ $D116$ により変調することにより、図15 (D)に示す

ような音質が改善されたオーディオデータD117を生成し、これを出力する。

因みに、図16はオーディオ信号処理装置100におけるクラス分類適応処理の処理手順を示し、オーディオ信号処理装置100はステップSP111から当該処理手順に入ると、続くステップSP112において入力オーディオデータD10の包絡線を包絡線算出部11において算出する。

この算出された包絡線は入力オーディオデータD10の特徴を表すものであり、オーディオ信号処理装置10は、ステップSP113に移ってクラス分類部14により包絡線に基づいてクラスを分類する。そしてオーディオ信号処理装置100はクラス分類の結果得られたクラスコードを用いて予測係数メモリ115から予測係数を読み出す。この予測係数は予め学習によりクラス毎に対応して格納されており、オーディオ信号処理装置100はクラスコードに対応した予測係数を読み出すことにより、このときの包絡線の特徴に合致した予測係数を用いることができる。

予測係数メモリ115から読み出された予測係数は、ステップSP114において包絡線予測演算部116の予測演算に用いられる。これにより、入力オーディオデータD10の包絡線の特徴に適応した予測演算により、所望とするオーディオデータD117を得るための新たな包絡線が算出される。ステップSP114において新たな包絡線が算出されると、オーディオ信号処理装置100は続くステップSP115において入力オーディオデータD10の搬送波を新たな包絡線により変調することにより、所望とするオーディオデータD117を得る。

かくして入力オーディオデータD10はその音質が改善されたオーディオデータD117に変換され、オーディオ信号処理装置100はステップSP116に移って当該処理手順を終了する。

次に、図14について上述した予測係数メモリ115に記憶するクラス毎の予測係数のセットを予め学習によって得るための学習回路について説明する。

図10との対応部分に同一符号を付して示す図16において、学習回路130は、高音質の教師オーディオデータD130を生徒信号生成フィルタ37に受け



る。生徒信号生成フィルタ 37 は、間引き率設定信号 D 39 により設定された間引き率で教師オーディオデータ D 130 を所定時間ごとに所定サンプル間引くようになされている。

この場合、生徒信号生成フィルタ 37 における間引き率によって、生成される予測係数が異なり、これに応じて上述のオーディオ信号処理装置 100 で再現されるオーディオデータも異なる。例えば、上述のオーディオ信号処理装置 100 においてサンプリング周波数を高くすることでオーディオデータの音質を向上しようとする場合、生徒信号生成フィルタ 37 ではサンプリング周波数を減らす間引き処理を行う。また、これに対して上述のオーディオ信号処理装置 100 において入力オーディオデータ D 130 の欠落したデータサンプルを補うことで音質の向上を図る場合には、これに応じて、生徒信号生成フィルタ 37 ではデータサンプルを欠落させる間引き処理を行うようになされている。

かくして、生徒信号生成フィルタ 37 は教師オーディオデータ D 130 から所定の間引き処理により生徒オーディオデータ D 37 を生成し、これを包絡線算出部 31 に供給する。

包絡線算出部 31 は生徒信号生成フィルタ 37 から供給された生徒オーディオデータ D 37 を所定時間毎の領域（この実施の形態の場合、例えば 6 サンプル毎とする）に分割した後、当該分割された各時間領域の波形について、図 4 について上述した包絡線算出方法によりその包絡線を算出する。

包絡線算出部 31 は生徒オーディオデータ D 37 のこのとき分割された時間領域の包絡線算出結果を生徒オーディオデータ D 37 の包絡線波形データ D 31 としてクラス分類部 34 に供給する。

クラス分類部 34 は、包絡線波形データ D 31 を圧縮して圧縮データパターンを生成する ADRC (Adaptive Dynamic Range Coding) 回路部と、包絡線波形データ D 31 の属するクラスコードを発生するクラスコード発生回路部とを有する。

ADRC 回路部は包絡線波形データ D 31 に対して、例えば 8 ビットから 2 ビ

ットに圧縮するような演算を行うことによりパターン圧縮データを形成する。このADRC回路部は、適応的量子化を行うものであり、ここでは、信号レベルの局所的なパターンを短い語長で効率的に表現することができるので、信号パターンのクラス分類のコード発生用に用いられる。

具体的には、包絡線波形上の6つの8ビットのデータ（包絡線波形データ）をクラス分類しようとする場合、 $2^{48}$ という膨大な数のクラスに分類しなければならない、回路上の負担が多くなる。そこで、この実施の形態のクラス分類部14ではその内部に設けられたADRC回路部で生成されるパターン圧縮データに基づいてクラス分類を行う。例えば6つの包絡線波形データに対して1ビットの量子化を実行すると、6つの包絡線波形データを6ビットで表すことができ、 $2^6 = 64$ クラスに分類することができる。

ここで、ADRC回路部は、切り出された領域内の包絡線のダイナミックレンジをDR、ビット割り当をm、各包絡線波形データのデータレベルをL、量子化コードをQとして、上述の(1)式と同様の演算により、領域内の最大値MAXと最小値MINとの間を指定されたビット長で均等に分割して量子化を行う。かくして包絡線算出部1において算出された包絡線上の6つの波形データが、それぞれ例えば8ビット ( $m=8$ ) で構成されているとすると、これらはADRC回路部においてそれぞれが2ビットに圧縮される。

このようにして圧縮された包絡線波形データをそれぞれ $q_n$  ( $n=1\sim6$ ) とすると、クラス分類部34に設けられたクラスコード発生回路部は、圧縮された包絡線波形データ $q_n$ に基づいて、上述の(2)式と同様の演算を実行することにより、そのブロック ( $q_1\sim q_6$ ) が属するクラスを示すクラスコードclassを算出し、当該算出されたクラスコードclassを表すクラスコードデータD34を予測係数算出部136に供給する。

このようにして、クラス分類部34は包絡線波形データD31のクラスコードデータD34を生成し、これを予測係数算出部136に供給する。また、予測係数算出部136には、生徒オーディオデータD37に基づいて算出された包絡線

波形データD31 ( $x_1, x_2, \dots, x_n$ ) が供給される。

予測係数算出部136は、クラス分類部34から供給されたクラスコードclassと、生徒オーディオデータD37に基づいて各クラスコードclass毎に算出された包絡線波形データD31と、入力端 $T_{IN}$ から供給された教師オーディオデータD130から包絡線算出部135において抽出された包絡線データ搬送波D135 (図15 (B)) とを用いて、正規方程式を立てる。

すなわち、生徒オーディオデータD37に基づいて算出された包絡線波形データD31のnサンプルのレベルをそれぞれ $x_1, x_2, \dots, x_n$ として、それぞれにpビットのADRCを行った結果の量子化データを $q_1, \dots, q_n$ とする。このとき、この領域のクラスコードclassを上述の(2)式のように定義する。そして、上述のように生徒オーディオデータD37に基づいて算出された包絡線波形データD31のレベルをそれぞれ、 $x_1, x_2, \dots, x_n$ とし、高音質の教師オーディオデータD130の包絡線波形のレベルをyとしたとき、クラスコード毎に、予測係数 $w_1, w_2, \dots, w_n$ によるnタップの線形推定式を設定する。これを上述の(4)式とする。学習前は、 $w_n$ が未定係数である。

学習回路130では、クラスコード毎に、複数のオーディオデータ(包絡線)に対して学習を行う。データサンプル数がMの場合、上述の(4)式に従って、上述の(5)式が設定される。但し $k=1, 2, \dots, M$ である。

$M > n$ の場合、予測係数 $w_1, \dots, w_n$ は一意的に決まらないので、誤差ベクトル $e$ の要素を(6)式によって定義し(但し、 $k=1, 2, \dots, M$ )、(7)式を最小にする予測係数を求める。いわゆる、最小自乗法による解法である。

ここで、(7)式による $w_n$ の偏微分係数を求める。この場合、(8)式を「0」にするように、各 $W_n$  ( $n=1 \sim 6$ ) を求めれば良い。

そして、(9)式及び(10)式のように、 $X_i, Y_i$ を定義すると、(8)式は行列を用いて(11)式として表される。

この方程式は、一般に正規方程式と呼ばれている。なお、ここでは $n=6$ である。

全ての学習用データ（教師オーディオデータD30、クラスコードclass、オーディオ波形データD33）の入力が完了した後、予測係数算出部36は各クラスコードclassに上述の(11)式に示した正規方程式を立てて、この正規方程式を掃き出し法等の一般的な行列解法を用いて、各 $W_u$ について解き、各クラスコード毎に、予測係数を算出する。予測係数算出部36は、算出された各予測係数(D36)を予測係数メモリ15に書き込む。

このような学習を行った結果、予測係数メモリ15には、量子化データ $q_1$ 、……、 $q_6$ で規定されるパターン毎に、高音質のオーディオデータ $y$ を推定するための予測係数が、各クラスコード毎に格納される。この予測係数メモリ15は、図14について上述したオーディオ信号処理装置100において用いられる。かかる処理により、線形推定式に従って通常のオーディオデータから高音質のオーディオデータを作成するための予測係数の学習が終了する。因みに、通常のオーディオデータから高音質のオーディオデータを作成する方法としては、線形推定式に限らず、種々の方法を適用し得る。

このように、学習回路130は、オーディオ信号処理装置100において補間処理を行う程度を考慮して、生徒信号生成フィルタ37で高音質の教師オーディオデータの間引き処理を行うことにより、オーディオ信号処理装置100における補間処理のための予測係数を生成することができる。

以上の構成において、オーディオ信号処理装置100は、包絡線算出部11において入力オーディオデータD10の時間波形領域での包絡線を算出する。この包絡線は入力オーディオデータD10の音質ごとに変わるもので、オーディオ信号処理装置100は入力オーディオデータD10の包絡線に基づいてそのクラスを特定する。

オーディオ信号処理装置10は、予め学習時に例えば歪みのない高音質のオーディオデータ（教師オーディオデータ）を得るための予測係数をクラス毎に求めておき、包絡線に基づいてクラス分類された入力オーディオデータD10の包絡線をそのクラスに応じた予測係数により予測演算する。これにより、入力オーデ

ィオデータD10の包絡線はその音質に応じた予測係数を用いて予測演算されるので、実用上十分な程度に音質が向上したオーディオデータ波形の包絡線が得られる。この包絡線に基づいて搬送波を変調することにより、音質が向上したオーディオデータが得られる。

また、クラス毎の予測係数を生成する学習時において、位相の異なる多数の教師オーディオデータについてそれぞれに対応した予測係数を求めておくことにより、オーディオ信号処理装置100における入力オーディオデータD10のクラス分類適応処理時に位相変動が生じても、位相変動に対応した処理を行うことができる。

以上の構成によれば、入力オーディオデータD10の時間波形領域における包絡線に基づいて入力オーディオデータD10をクラス分類し、当該クラス分類された結果に基づく予測係数を用いて入力オーディオデータD10の包絡線を予測演算するようにしたことにより、入力オーディオデータD10を一段と高音質のオーディオデータD117に変換し得る包絡線を生成できる。

また上述の実施の形態においては、包絡線データD11に基づいてクラス分類する場合について述べたが、本発明はこれに限らず、クラス分類部14に入力オーディオデータD10を入力し、当該クラス分類部14において入力オーディオデータD10の波形に基づくクラス分類を行うと共に、包絡線算出部11において包絡線のクラス分類を行い、クラス分類部14においてこれら2つのクラスを統合することにより、入力オーディオデータD10の波形とその包絡線の両方に基づくクラス分類を行うようにしても良い。

### (3) 他の実施の形態

なお上述の実施の形態においては、図5について上述した包絡線算出方法を用いる場合について述べたが、本発明はこれに限らず、例えば単にピーク値を結ぶ方法等、他の種々の包絡線算出方法を適用することができる。

また上述の実施の形態においては、予測方式として線形一次による手法を用いる場合について述べたが、本発明はこれに限らず、要は学習した結果を用いるよ

うにすれば良く、例えば多次関数による手法、さらには入力端子 $T_{IN}$ から供給されるデジタルデータが画像データの場合には、画素値自体から予測する手法等、種々の予測方式を適用することができる。

また上述の実施の形態においては、クラス分類部14においてADRCにより圧縮データパターンを生成する場合について述べたが、本発明はこれに限らず、可逆符号化(DPCM: Differential Pulse Code Modulation)又はベクトル量子化(VQ: Vector Quantize)等の圧縮手段を用いるようにしても良い。

また上述の実施の形態においては、学習回路30の生徒信号生成フィルタ37において所定サンプル数を間引く場合について述べたが、本発明はこれに限らず、例えばビット数を削減する等、他の種々の方法を適用することができる。

また上述の実施の形態においては、オーディオデータを処理する装置に本発明を適用する場合について述べたが、本発明はこれに限らず、画像データや他の種々のデータを変換する場合に広く適用することができる。

上述のように本発明によれば、入力デジタル信号の包絡線に基づいて入力デジタル信号のクラスを分類し、当該分類されたクラスに対応した予測方式で入力デジタル信号を変換するようにしたことにより、一段と入力デジタル信号の特徴に適応した変換を行うことができる。

#### 産業上の利用の可能性

本発明は、デジタル信号に対してデータの補間処理を行うレートコンバータ、PCM復号装置やオーディオ信号処理装置に利用し得る。

## 請 求 の 範 囲

1. 入力デジタル信号を変換するデジタル信号処理装置において、  
上記入力デジタル信号の包絡線を算出する包絡線算出手段と、  
上記算出された包絡線に基づいて上記入力デジタル信号のクラスを分類する  
クラス分類手段と、  
上記分類されたクラスに対応した予測方式で上記入力デジタル信号を予測演  
算することにより上記入力デジタル信号を変換してなるデジタル信号を生成  
する予測演算手段と  
を具えることを特徴とするデジタル信号処理装置。
2. 上記入力デジタル信号はデジタルオーディオ信号である  
ことを特徴とする請求の範囲第1項に記載のデジタル信号処理装置。
3. 上記予測演算手段は、予め所望とするデジタル信号に基づいて学習によ  
り生成されている予測係数を用いる  
ことを特徴とする請求の範囲第1項に記載のデジタル信号処理装置。
4. 入力デジタル信号を変換するデジタル信号処理方法において、  
上記入力デジタル信号の包絡線を算出する包絡線算出ステップと、  
上記算出された包絡線に基づいて上記入力デジタル信号のクラスを分類する  
クラス分類ステップと、  
上記分類されたクラスに対応した予測方式で上記入力デジタル信号を予測演  
算することにより上記入力デジタル信号を変換してなるデジタル信号を生成  
する予測演算ステップと  
を具えることを特徴とするデジタル信号処理方法。

5. 上記入力デジタル信号はデジタルオーディオ信号であることを特徴とする請求の範囲第4項に記載のデジタル信号処理方法。
6. 上記予測演算ステップでは、予め所望とするデジタル信号に基づいて学習により生成されている予測係数が用いられることを特徴とする請求の範囲第4項に記載のデジタル信号処理方法。
7. 入力デジタル信号を変換するデジタル信号処理装置の上記変換処理の予測演算に用いられる予測係数を生成する学習装置において、  
所望とするデジタル信号から当該デジタル信号を劣化させた生徒デジタル信号を生成する生徒デジタル信号生成手段と、  
上記生徒デジタル信号の包絡線を算出する包絡線算出手段と、  
上記算出された包絡線に基づいて上記生徒デジタル信号のクラスを分類するクラス分類手段と、  
上記入力デジタル信号と上記生徒デジタル信号とに基づいて上記クラスに対応する予測係数を算出する予測係数算出手段と  
を具えることを特徴とする学習装置。
8. 上記入力デジタル信号はデジタルオーディオ信号であることを特徴とする請求の範囲第7項に記載の学習装置。
9. 入力デジタル信号を変換するデジタル信号処理装置の上記変換処理の予測演算に用いられる予測係数を生成する学習方法において、  
所望とするデジタル信号から当該デジタル信号を劣化させた生徒デジタル信号を生成する生徒デジタル信号生成ステップと、  
上記生徒デジタル信号の包絡線を算出する包絡線算出ステップと、  
上記算出された包絡線に基づいて上記生徒デジタル信号のクラスを分類する



クラス分類ステップと、

上記入力デジタル信号と上記生徒デジタル信号とに基づいて上記クラスに対応する予測係数を算出する予測係数算出ステップと  
を具えることを特徴とする学習方法。

10. 上記入力デジタル信号はデジタルオーディオ信号であることを特徴とする請求の範囲第9項に記載の学習方法。

11. 入力デジタル信号を変換するデジタル信号処理装置において、  
上記入力デジタル信号の包絡線を算出する包絡線算出手段と、  
上記算出された包絡線に基づいて上記デジタル信号のクラスを分類するクラス分類手段と、  
上記分類されたクラスに対応した予測方式で新たな包絡線を算出する包絡線予測演算手段と、  
上記入力デジタル信号から搬送波を抽出する搬送波抽出手段と、  
上記包絡線予測演算手段によって算出された上記新たな包絡線に基づいて上記搬送波を変調することにより上記入力デジタル信号を変換してなる新たなデジタル信号を生成する変調手段と  
を具えることを特徴とするデジタル信号処理装置。

12. 上記入力デジタル信号はデジタルオーディオ信号であることを特徴とする請求の範囲第11項に記載のデジタル信号処理装置。

13. 上記包絡線予測演算手段は、予め所望とするデジタル信号に基づいて学習により生成されている予測係数を用いる  
ことを特徴とする請求の範囲第11項に記載のデジタル信号処理装置。

14. 入力デジタル信号を変換するデジタル信号処理方法において、  
上記入力デジタル信号の包絡線を算出する包絡線算出ステップと、  
上記算出された包絡線に基づいて上記デジタル信号のクラスを分類するクラス分類ステップと、  
上記分類されたクラスに対応した予測方式で新たな包絡線を算出する包絡線予測演算ステップと、  
上記入力デジタル信号から搬送波を抽出するステップと、  
上記包絡線予測演算ステップによって算出された上記新たな包絡線に基づいて上記搬送波を変調することにより上記入力デジタル信号を変換してなる新たなデジタル信号を生成するステップと  
を具えることを特徴とするデジタル信号処理方法。

15. 上記入力デジタル信号はデジタルオーディオ信号であることを特徴とする請求の範囲第14項に記載のデジタル信号処理方法。

16. 上記包絡線予測演算ステップでは、予め所望とするデジタル信号に基づいて学習により生成されている予測係数が用いられる  
ことを特徴とする請求の範囲第14項に記載のデジタル信号処理方法。

17. 入力デジタル信号を変換するデジタル信号処理装置の上記変換処理の予測演算に用いられる予測係数を生成する学習装置において、  
所望とするデジタル信号から当該デジタル信号を劣化させた生徒デジタル信号を生成する生徒デジタル信号生成手段と、  
上記生徒デジタル信号の包絡線を算出する第1の包絡線算出手段と、  
上記算出された包絡線に基づいて上記生徒デジタル信号のクラスを分類するクラス分類手段と、  
上記入力デジタル信号の包絡線を算出する第2の包絡線算出手段と、

上記第 1 の包絡線算出手段により算出された上記生徒デジタル信号の包絡線と上記第 2 の包絡線算出手段により算出された上記入力デジタル信号の包絡線とに基づいて上記クラスに対応する予測係数を算出する予測係数算出手段とを具えることを特徴とする学習装置。

18. 上記入力デジタル信号はデジタルオーディオ信号であることを特徴とする請求の範囲第 17 項に記載の学習装置。

19. 入力デジタル信号を変換するデジタル信号処理装置の上記変換処理の予測演算に用いられる予測係数を生成する学習方法において、

所望とするデジタル信号から当該デジタル信号を劣化させた生徒デジタル信号を生成する生徒デジタル信号生成ステップと、

上記生徒デジタル信号の包絡線を算出する第 1 の包絡線算出ステップと、

上記算出された包絡線に基づいて上記生徒デジタル信号のクラスを分類するクラス分類ステップと、

上記入力デジタル信号の包絡線を算出する第 2 の包絡線算出ステップと、

上記算出された上記生徒デジタル信号の包絡線と上記算出された上記入力デジタル信号の包絡線とに基づいて上記クラスに対応する予測係数を算出する予測係数算出ステップと

を具えることを特徴とする学習方法。

20. 上記入力デジタル信号はデジタルオーディオ信号であることを特徴とする請求の範囲第 19 項に記載の学習方法。

21. 入力デジタル信号の包絡線を算出する包絡線算出ステップと、

上記算出された包絡線に基づいて上記入力デジタル信号のクラスを分類するクラス分類ステップと、

上記分類されたクラスに対応した予測方式で上記入力デジタル信号を予測演算することにより上記入力デジタル信号を変換してなるデジタル信号を生成する予測演算ステップと

を含むプログラムをデジタル信号処理装置に実行させるプログラム格納媒体

22. 所望とするデジタル信号から当該デジタル信号を劣化させた生徒デジタル信号を生成する生徒デジタル信号生成ステップと、

上記生徒デジタル信号の包絡線を算出する包絡線算出ステップと、

上記算出された包絡線に基づいて上記生徒デジタル信号のクラス进行分类するクラス分類ステップと、

上記デジタル信号と上記生徒デジタル信号とに基づいて上記クラスに対応する予測係数を算出する予測係数算出ステップと

を含むプログラムを学習装置に実行させるプログラム格納媒体。

23. 入力デジタル信号の包絡線を算出する包絡線算出ステップと、

上記算出された包絡線に基づいて上記デジタル信号のクラス进行分类するクラス分類ステップと、

上記分類されたクラスに対応した予測方式で新たな包絡線を算出する包絡線予測演算ステップと、

上記入力デジタル信号から搬送波を抽出する搬送波抽出ステップと、

上記包絡線予測演算手段によって算出された上記新たな包絡線に基づいて上記搬送波を変調することにより上記入力デジタル信号を変換してなる新たなデジタル信号を生成する変調ステップと

を含むプログラムをデジタル信号処理装置に実行させるプログラム格納媒体

24. 所望とするデジタル信号から当該デジタル信号を劣化させた生徒デジタル信号を生成する生徒デジタル信号生成ステップと、

上記生徒デジタル信号の包絡線を算出する包絡線算出ステップと、

上記算出された包絡線に基づいて上記生徒デジタル信号のクラスを分類するクラス分類ステップと、

上記デジタル信号の包絡線を算出する包絡線算出ステップと、

上記算出された上記生徒デジタル信号の包絡線と上記算出された上記デジタル信号の包絡線とに基づいて上記クラスに対応する予測係数を算出する予測係数算出ステップと

を含むプログラムを学習装置に実行させるプログラム格納媒体。

**THIS PAGE BLANK (USPTO)**

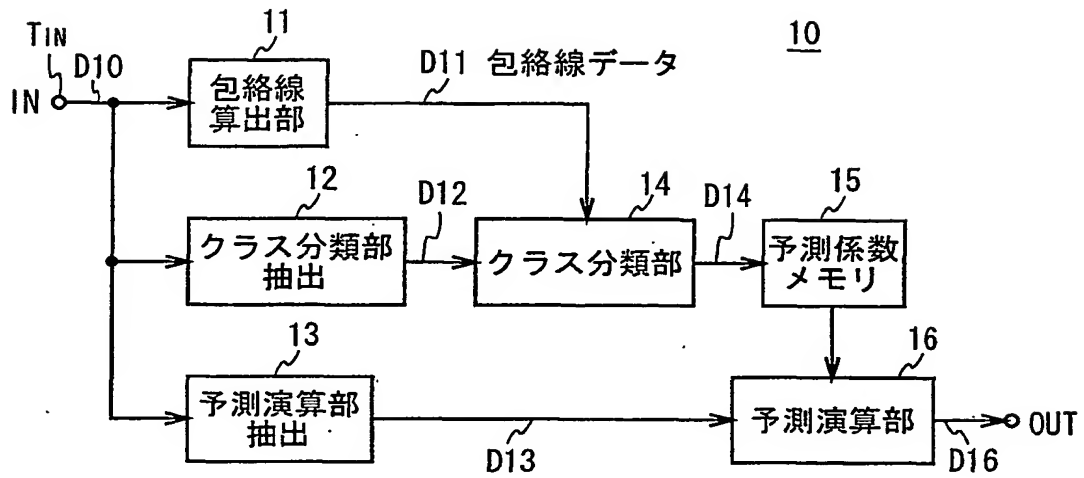


図 1

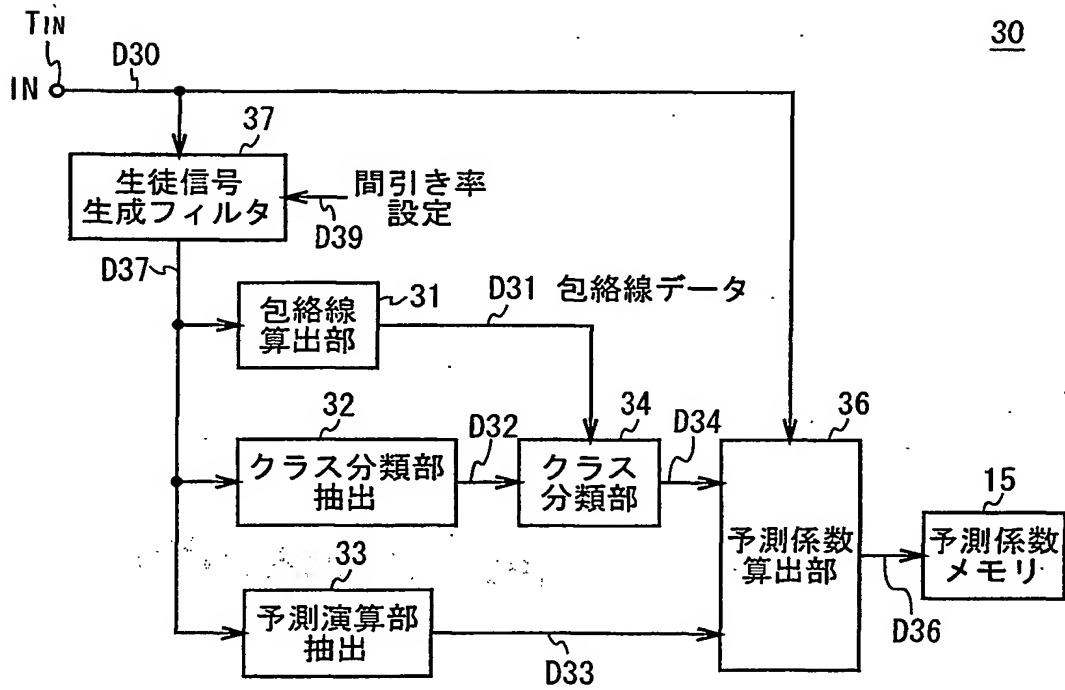


図 1 1

**THIS PAGE BLANK (USPTO)**



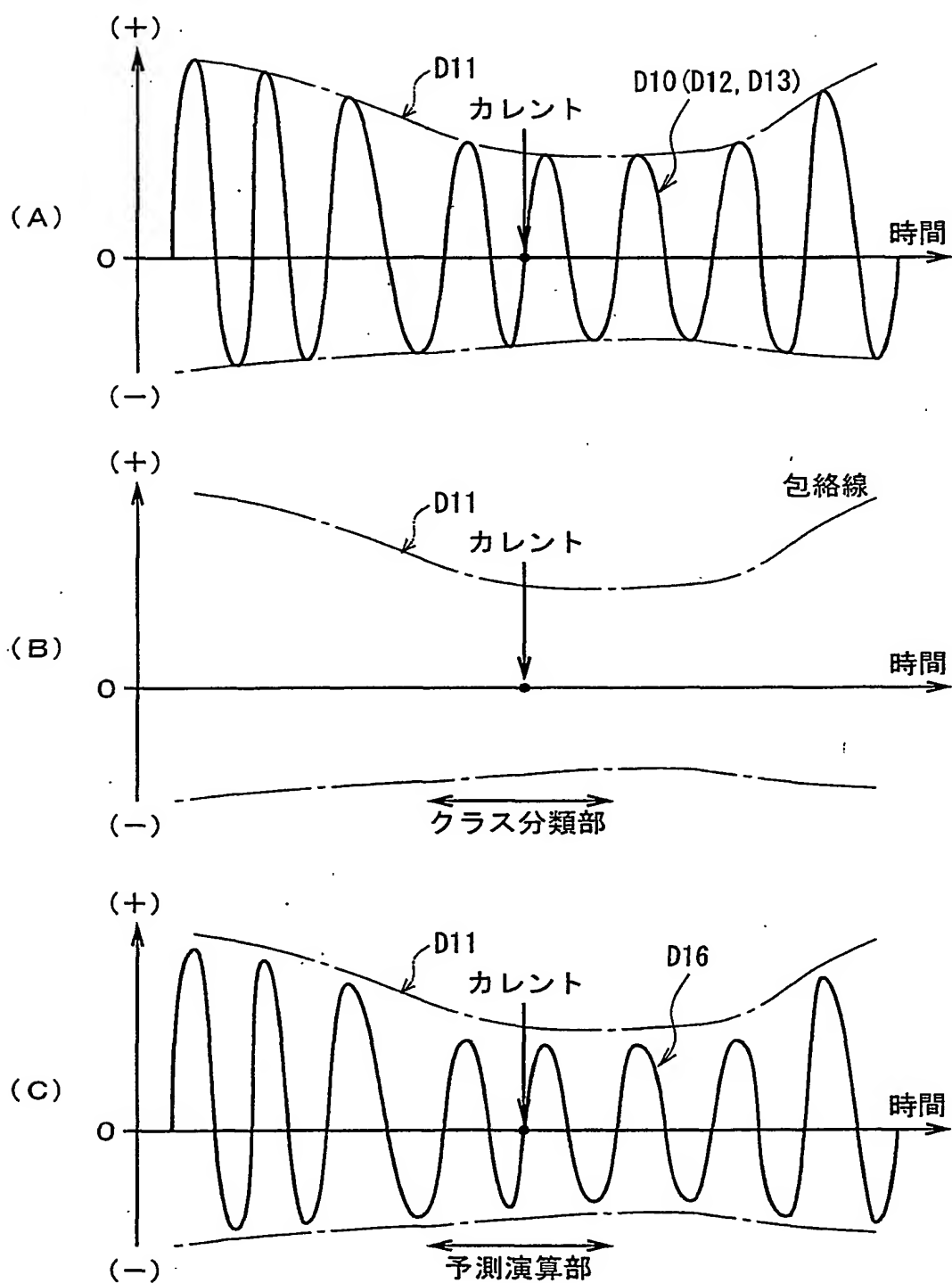


図 2

**THIS PAGE BLANK (USPTO)**

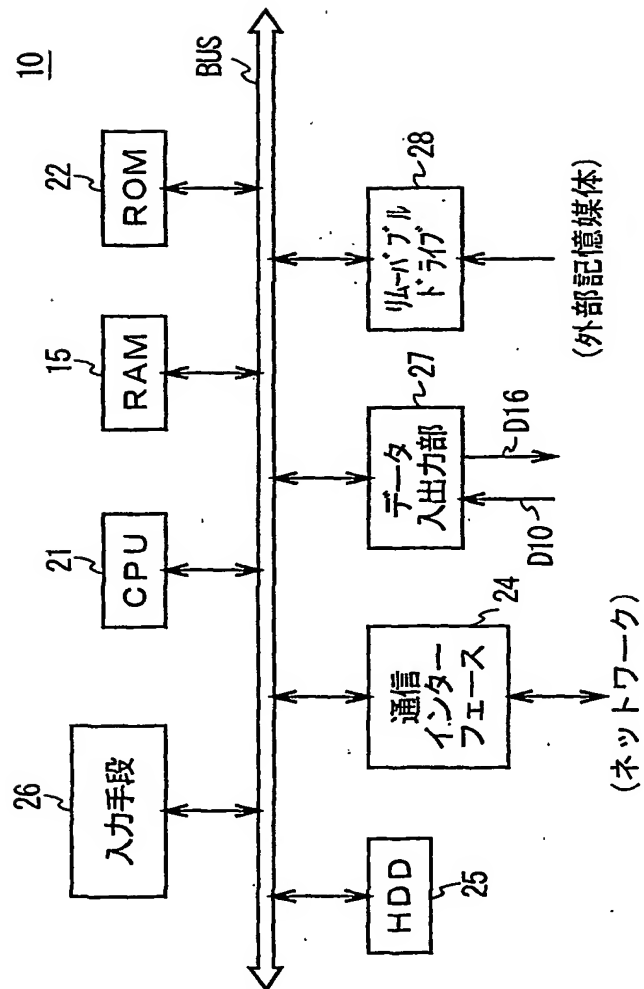


図 3

**THIS PAGE BLANK (USPTO)**

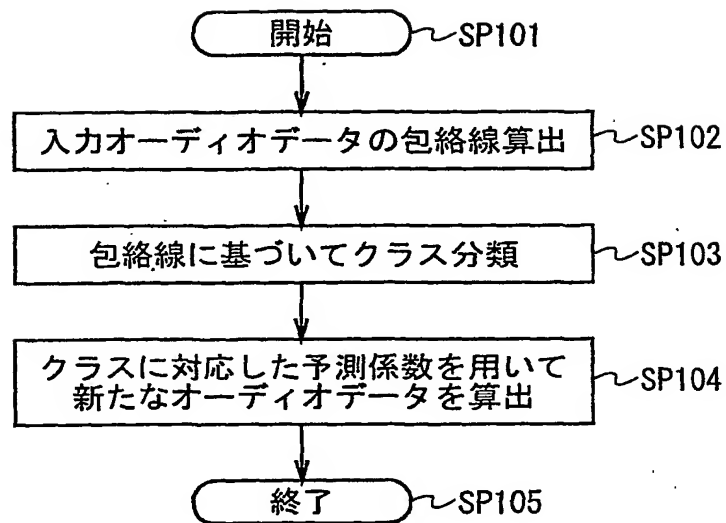


図 4

**THIS PAGE BLANK (USPTO)**

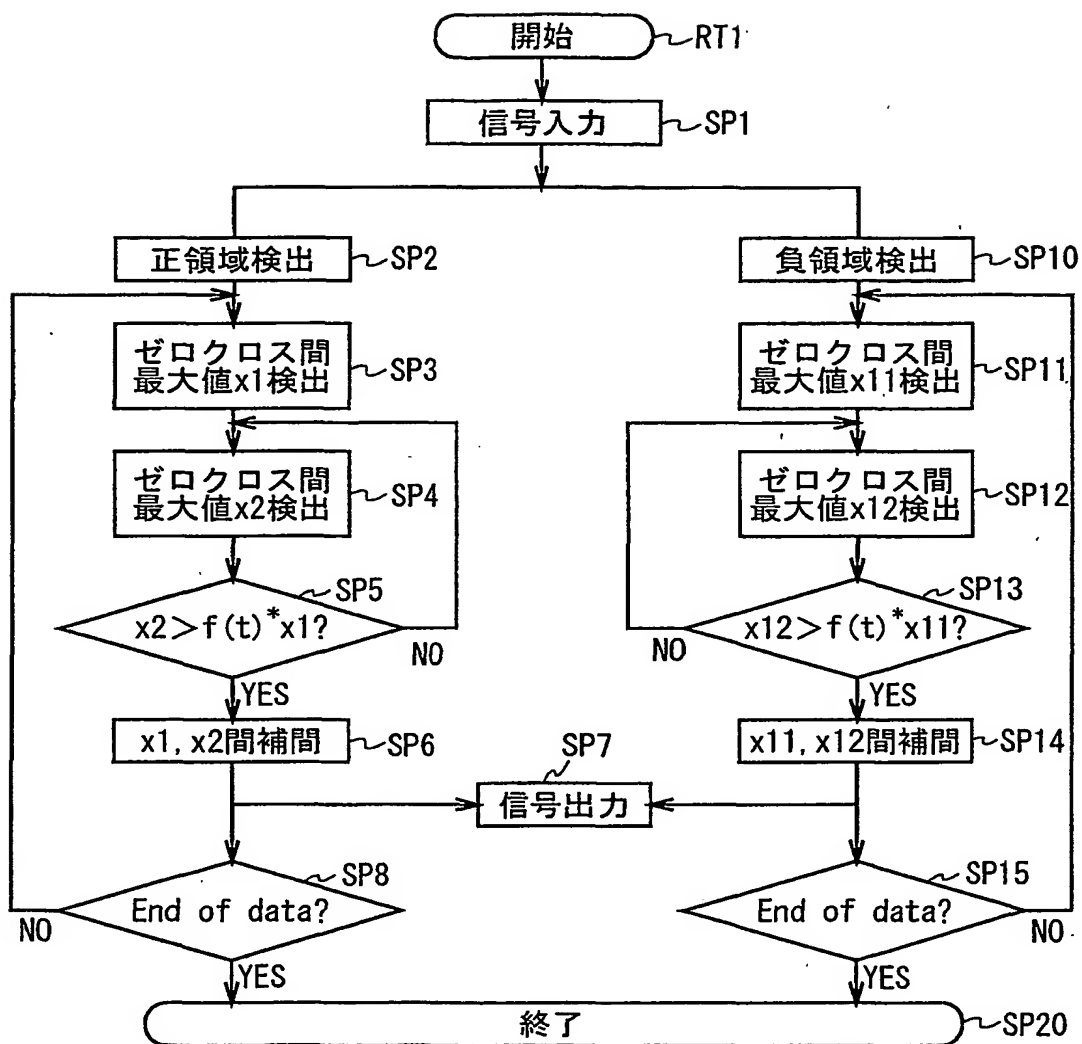


図 5

**THIS PAGE BLANK (USPTO)**



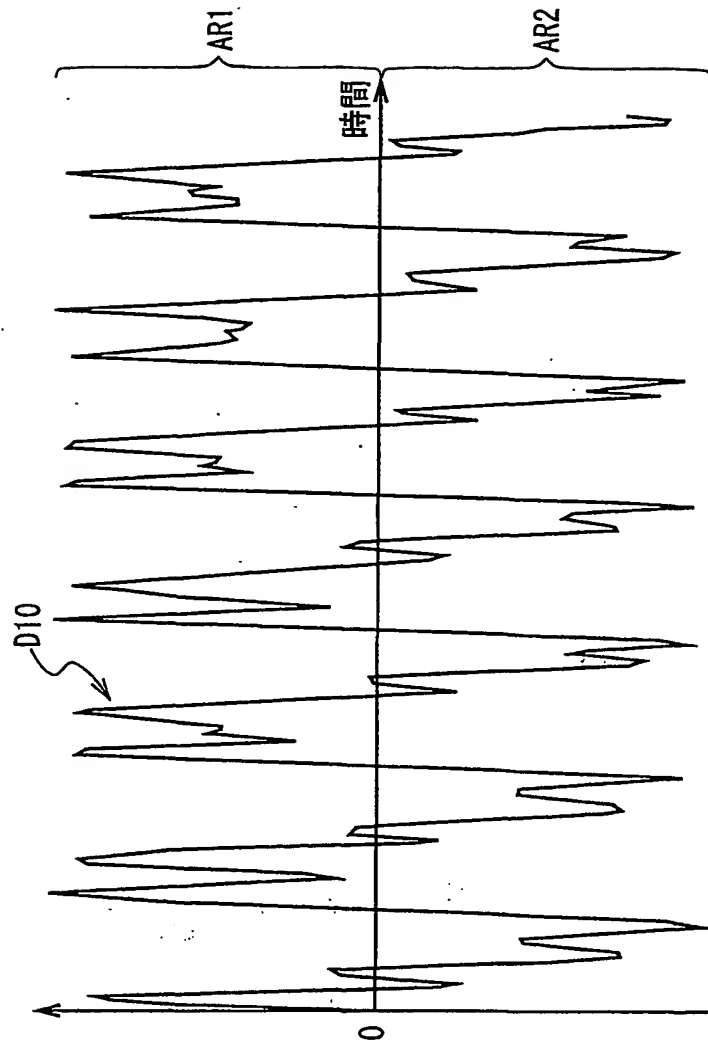


図 6

**THIS PAGE BLANK (USPTO)**

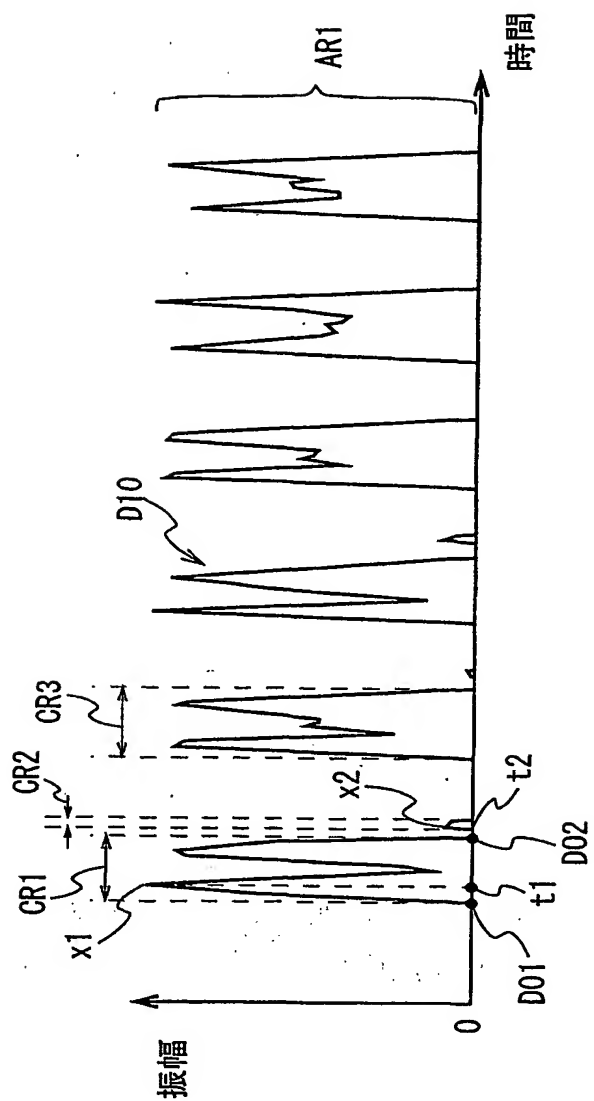


図 7

**THIS PAGE BLANK (USPTO)**

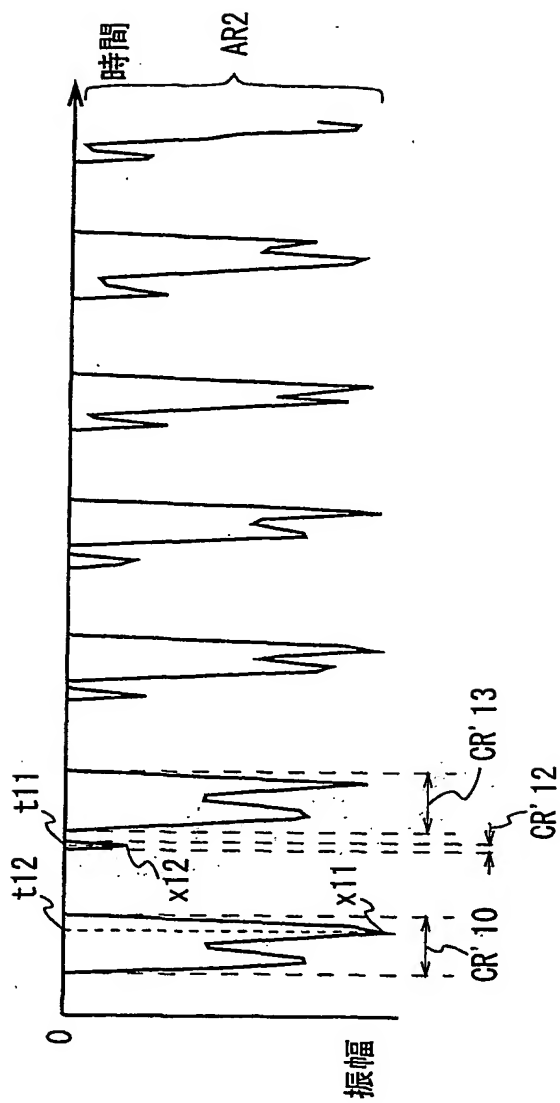


図 8

**THIS PAGE BLANK (11/27/71)**

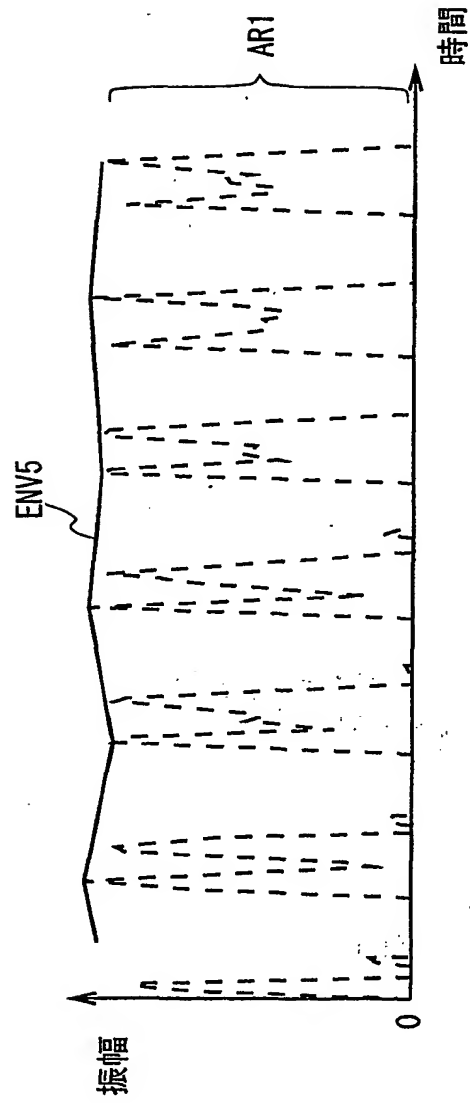


図 9

**THIS PAGE BLANK (USPTO)**



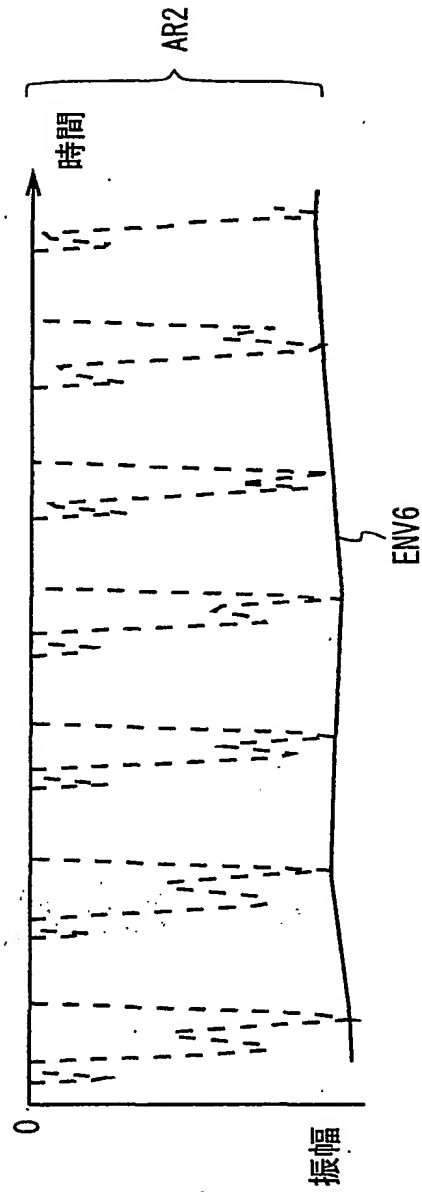


図 10

**THIS PAGE BLANK (USPTO)**

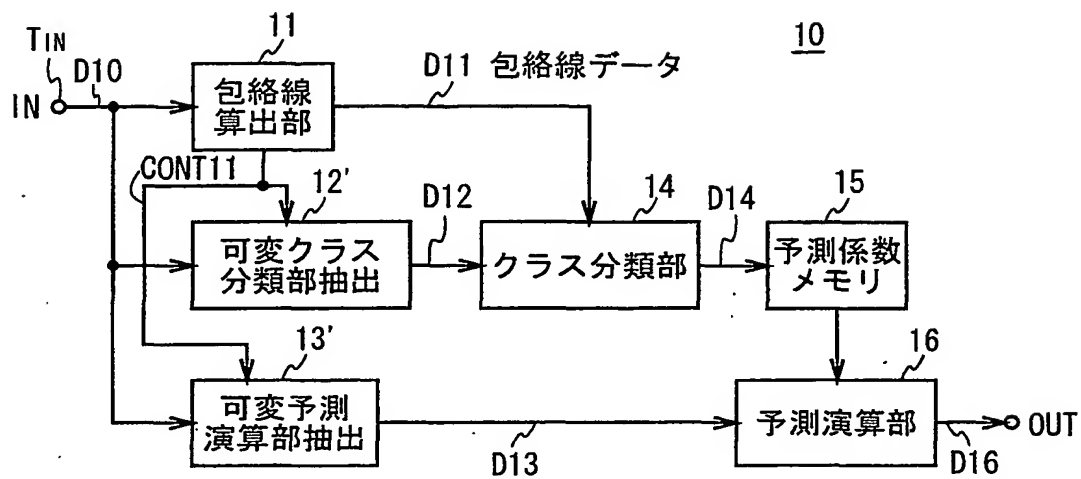


図 1 2

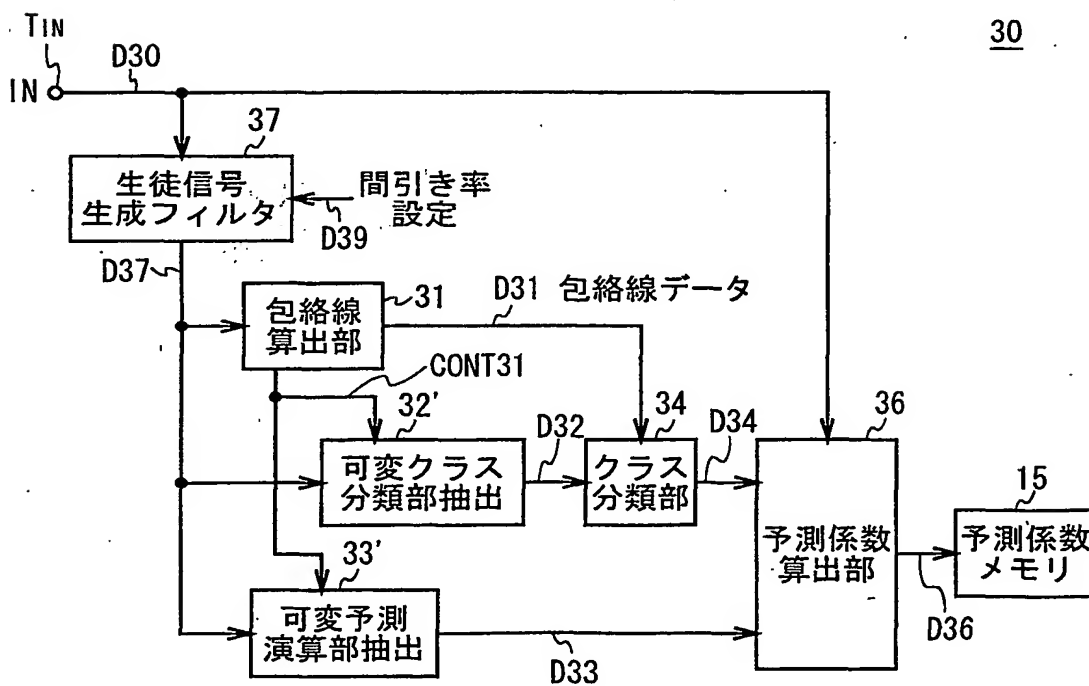


図 1 3

**THIS PAGE BLANK (USPTO)**

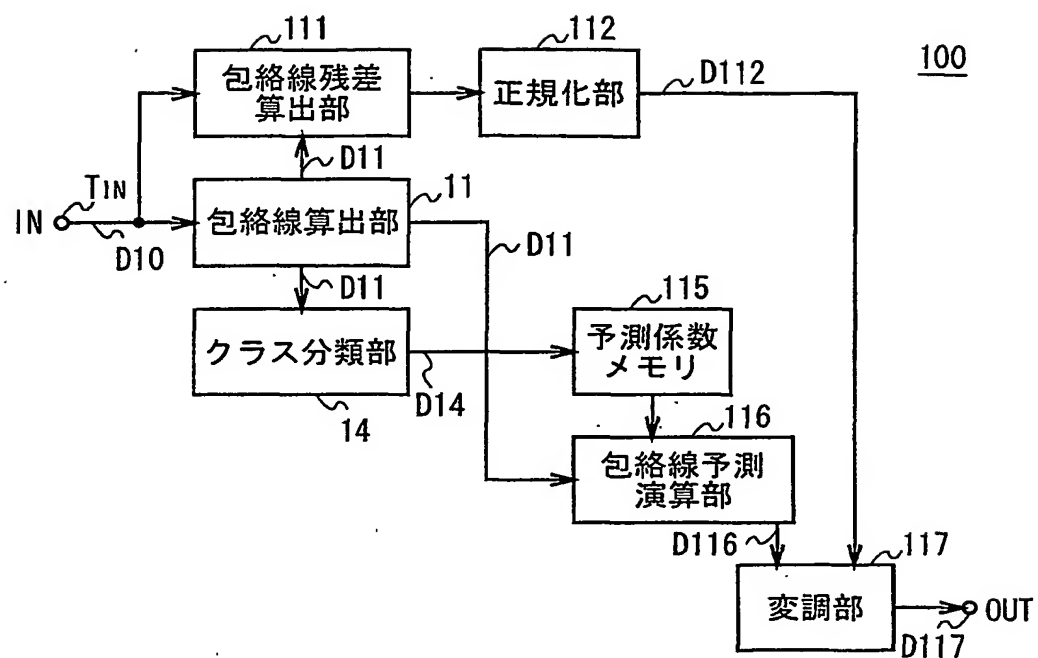


図 1 4

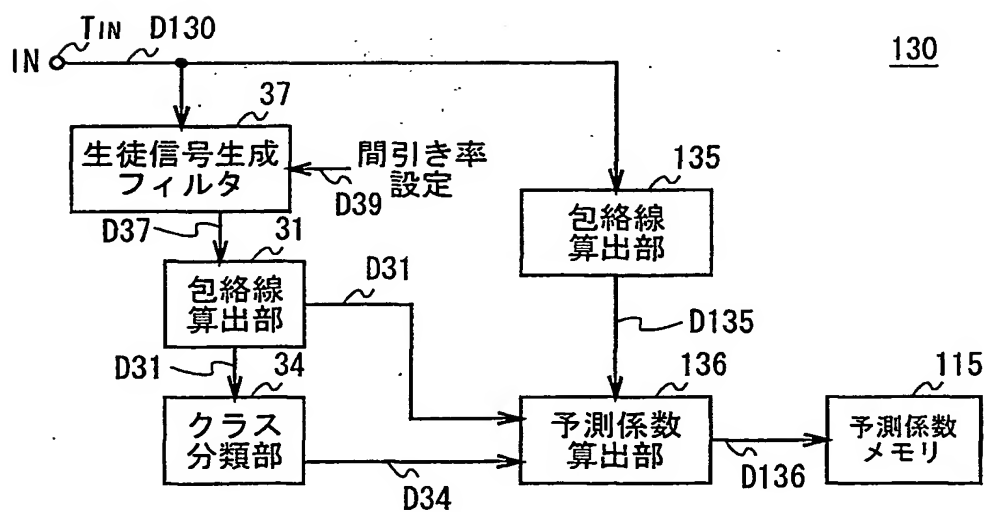


図 1 7

**THIS PAGE BLANK (USPTO)**

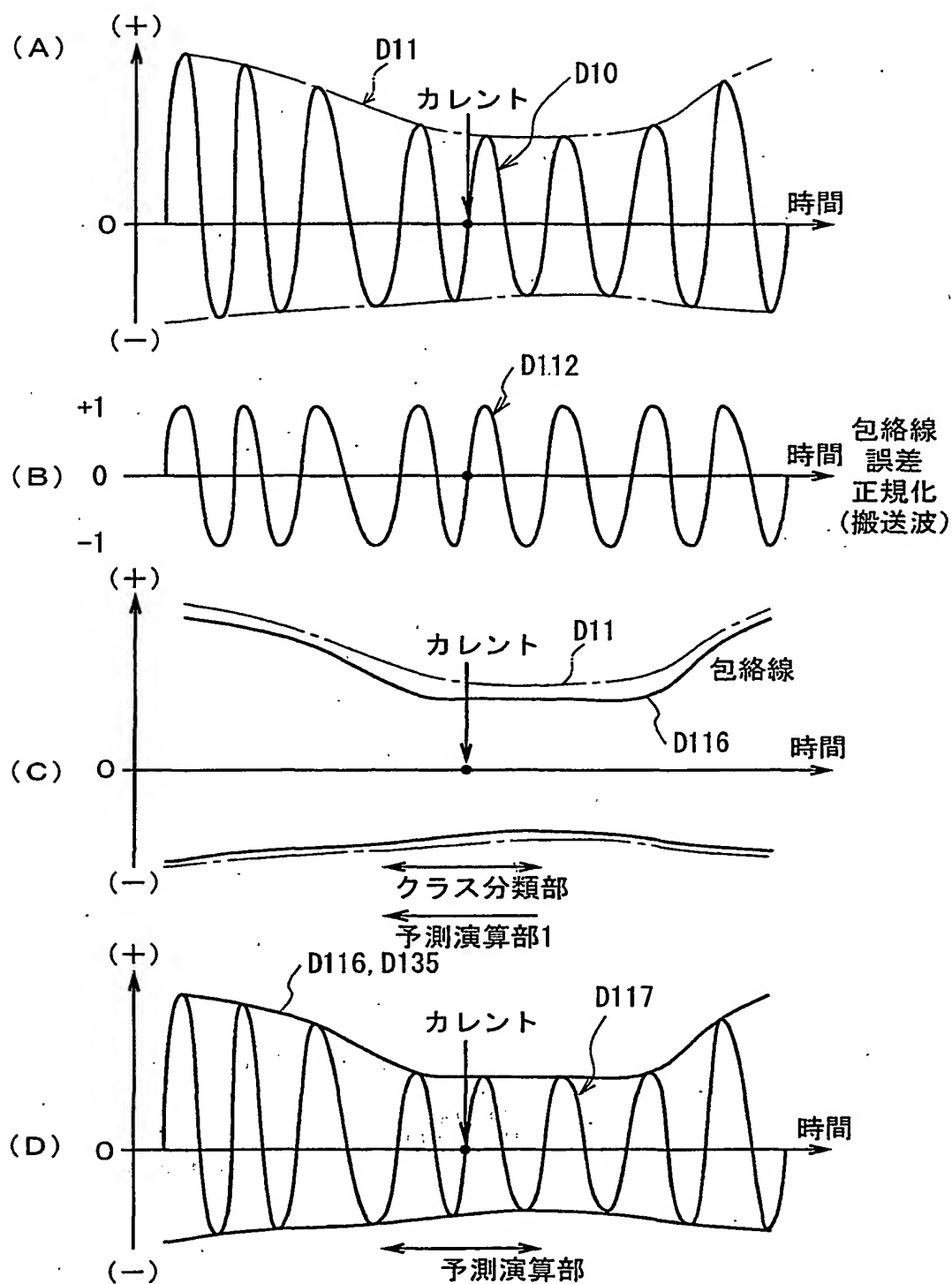


図 1 5

**THIS PAGE BLANK (USPTO)**



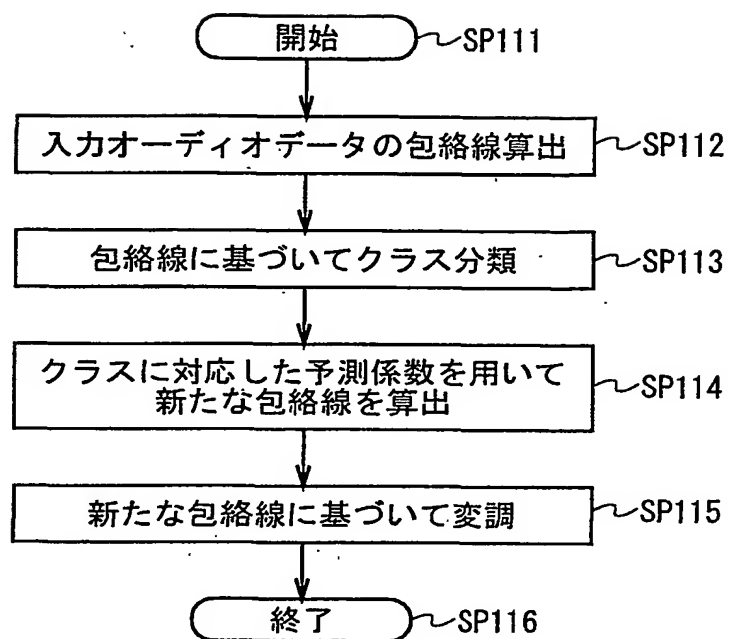


図 16

**THIS PAGE BLANK (USPTO)**

## 符 号 の 説 明

10、100……オーディオ信号処理装置、11、31……包絡線算出部、14、34……クラス分類部、15、115……予測係数メモリ、16……予測演算部、36……予測係数算出部、37……生徒信号生成フィルタ、111……包絡線残差算出部、112……正規化部、116……包絡線予測演算部、117…変調部

**THIS PAGE BLANK (USPTO)**

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP01/06593

A. CLASSIFICATION OF SUBJECT MATTER Int.Cl <sup>7</sup> G10L13/00, H03M7/32-38, H03H17/00-06		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) Int.Cl <sup>7</sup> G10L13/00, H03M3/00-04, H03M7/32-38, H03H17/00-06, H04N7/00		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Jitsuyo Shinan Koho 1922-1996 Toroku Jitsuyo Shinan Koho 1971-2001 Kokai Jitsuyo Shinan Koho 1971-2001 Jitsuyo Shinan Toroku Koho 1996-2001		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 04-115628 A (Sony Corporation), 16 April, 1992 (16.04.92), Full text; Figs. 1 to 8 (Family: none)	1-24
A	JP 05-297898 A (Sony Corporation), 12 November, 1993 (12.11.93), Full text; Figs. 1 to 4 (Family: none)	1-24
A	JP 05-323999 A (Kokusai Electric Co., Ltd.), 07 December, 1993 (07.12.93), Full text; Figs. 1 to 3 (Family: none)	1-24
A	JP 06-051800 A (Sony Corporation), 25 February, 1994 (25.02.94), Full text; Figs. 1 to 9 (Family: none)	1-24
A	JP 10-313251 A (Sony Corporation), 24 November, 1998 (24.11.98), Full text; Figs. 1 to 7 (Family: none)	1-24
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family	
Date of the actual completion of the international search 23 October, 2001 (23.10.01)		Date of mailing of the international search report 13 November, 2001 (13.11.01)
Name and mailing address of the ISA/ Japanese Patent Office		Authorized officer
Facsimile No.		Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP01/06593

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 11-027564 A (Sony Corporation), 29 January, 1999 (29.01.99), Full text; Figs. 1 to 12 & EP 912045 A1 & WO 98/051072 A1	1-24
A	JP 2000-32402 A (Sony Corporation), 28 January, 2000 (28.01.00), Full text; Figs. 1 to 8 (Family: none)	1-24
A	JP 2000-78534 A (Sony Corporation), 14 March, 2000 (14.03.00), Full text; Figs. 1 to 16 (Family: none)	1-24
A	JP 57-144600 A (NEC Corporation), 07 September, 1982 (07.09.82), Full text; Figs. 1 to 2 (Family: none)	1-24
A	JP 60-195600 A (Sanyo Electric Co., Ltd.), 04 October, 1985 (04.10.85), Full text; Figs. 1 to 2 (Family: none)	1-24
A	US 5903866 A (Lucent Technologies Inc.), 10 March, 1997 (10.03.97), Full text; Figs. 1 to 6 & JP 10-307599 A & EP 865028 A1	1-24

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl<sup>1</sup> G10L13/00 H03M7/32-38 H03H17/00-06

## B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl<sup>1</sup> G10L13/00 H03M3/00-04 H03M7/32-38  
H03H17/00-06 H04N7/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2001年
日本国登録実用新案公報	1971-2001年
日本国実用新案登録公報	1996-2001年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	J P 04-115628 A (ソニー株式会社) 16.4月.1992 (16.04.92) 全文, 第1-8図 (ファミリーなし)	1-24
A	J P 05-297898 A (ソニー株式会社) 12.11月.1993 (12.11.93) 全文, 第1-4図 (ファミリーなし)	1-24

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの  
「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの  
「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)  
「O」 口頭による開示、使用、展示等に言及する文献  
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&amp;」 同一パテントファミリー文献

国際調査を完了した日

23.10.01

国際調査報告の発送日

13.11.01

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

南 義明

5C

9381

電話番号 03-3581-1101 内線 3541

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 05-323999 A (国際電気株式会社) 7.12月.1993 (07.12.93) 全文, 第1-3図 (ファミリーなし)	1-24
A	JP 06-051800 A (ソニー株式会社) 25.2月.1994 (25.02.94) 全文, 第1-9図 (ファミリーなし)	1-24
A	JP 10-313251 A (ソニー株式会社) 24.11月.1998 (24.11.98) 全文, 第1-7図 (ファミリーなし)	1-24
A	JP 11-027564 A (ソニー株式会社) 29.1月.1999 (29.01.99) 全文, 第1-12図 & EP 912045 A1 & WO 98/051072 A1	1-24
A	JP 2000-32402 A (ソニー株式会社) 28.1月.2000 (28.01.00) 全文, 第1-8図 (ファミリーなし)	1-24
A	JP 2000-78534 A (ソニー株式会社) 14.3月.2000 (14.03.00) 全文, 第1-16図 (ファミリーなし)	1-24
A	JP 57-144600 A (日本電気株式会社) 7.9月.1982 (07.09.82) 全文, 第1-2図 (ファミリーなし)	1-24
A	JP 60-195600 A (三洋電機株式会社) 4.10月.1985 (04.10.85) 全文, 第1-2図 (ファミリーなし)	1-24
A	US 5903866 A (Lucent Technologies Inc.) 10.3月.1997 (10.03.97) 全文, 第1-6図 & JP 10-307599 A & EP 865028 A1	1-24